FACULTY WORKING PAPER 92-0113

A Study of Comparative Judgments Based on Numerical and Verbal Attribute Labels

Madhubalan Viswanathan
Department of Business Adminstration

University of Illinois

Sunder Narayanan
Departments of Advertising and Business Administration
University of Illinois

Bureau of Economic and Business Research
College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 92-0113

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

March 1992

A Study of Comparative Judgments Based on Numerical and Verbal Attribute Labels

Madhubalan Viswanathan*
Sunder Narayanan

* Please address comments to Madhubalan Viswanathan, Department of Business Administration, University of Illinois, Champaign, IL 61820; Tel. (217) 333-4550

A Study of Comparative Judgments Based on Numerical and Verbal Attribute Labels

Abstract

Consumers often compare brands on a specific attribute using the information available to them. The paper examines comparative judgments of brands based on numerical and verbal labels. Using a comparative judgment task from cognitive psychology, comparisons of pairs of numerical and/or verbal labels are studied in two experiments. Results of the first experiment suggest that comparing a pair of numerical labels may be easier than a pair of verbal labels or a numerical-verbal pair. Further, several past findings are replicated in the context of brand attribute comparisons. In addition, results also suggest that the utility properties of the attribute may provide a reference point for the comparisons of numerical and verbal labels. A second study replicates the findings of the first study and provides support for the effect of the utility property of an attribute.

A consumer looking to buy her favorite brand of breakfast cereal searches the supermarket shelf to locate the product. While looking for the product, she spots a new brand of cereal on the shelf. She picks up a box of the new cereal to compare with her favorite brand. Being a health conscious consumer, she compares the two brands on several attributes such as calorie content, fiber content and sodium content. She notices on the package that the new brand has 100 calories, 4 grams of dietary fiber and 200 grams of sodium per serving, while her regular brand has 150 calories, 3 grams of dietary fiber and 200 grams of sodium per serving. The consumer is particularly fond of the crunchy taste of her regular brand. The new brand does not make any claims of crunchiness on the box. Due to the absence of information on the crunchiness of the new brand, she buys her regular brand. Later that evening, while reading the latest issue of Consumer Reports, she finds a comparison of various cereals on several attributes. She notices that her favorite brand gets a "much better than average" rating on crunchiness, while the new brand in the store gets a "better than average" rating. She tries to recall the information she obtained earlier in the store about the new brand. She recalls that the new brand had lower calories than her regular brand, though she forgot the precise numerical information. She also recalls that the new brand had a dietary fiber content of 4 grams compared to 3 grams for her regular brand. She consider all the information, including the information on crunchiness from Consumer Reports,

and decides to try the new brand next time.

In typical consumer settings, information about product attributes are often conveyed using numerical labels (such as "200" calories) or verbal labels (such as "much better than average" crunchiness). In recent times, several researchers in consumer behavior have focussed on differences between numerical and verbal information (cf. Viswanathan and Narayanan, 1992; Viswanathan and Childers, 1992). Since an important element of consumer decision making involves comparing brands across attributes, an understanding of how consumers compare verbal and numerical labels is an important issue in consumer research. In the example presented earlier, the consumer compared the two brands using information available in both numerical and verbal modes. She found that on the attribute fiber content, numerical information was available. On another attribute calories, numerical information was available for one brand, though she could only recall verbal information for the other brand. On the attribute crunchiness, only verbal information was available. Consumers typically face this type of situations where they have to compare brands on attributes, using numerical and/or verbal labels.

This study addresses the issue of how consumers compare brands using attributes with numerical and verbal labels. Specifically, it focuses on comparisons of brands on attributes when information is presented using verbal and/or numerical labels. First, relevant research from psychology on comparative

judgments is briefly reviewed. This is followed by a discussion of comparative judgments of brand attributes and some hypotheses. Finally, the details of two experiments conducted to test the hypotheses are presented.

LITERATURE REVIEW

Research in psychology has studied comparative judgments of stimuli along various dimensions (Banks et al. 1976; Holyoak 1978; Jaffe-Katz et al. 1989; Moyer and Landauer 1967). The comparative judgment task requires individuals to compare stimuli across a dimension and make judgments about the magnitude of the stimuli along the dimension. For example, subjects may be required to identify the larger of two stimuli (such as an elephant and a mouse) along a dimension such as size. Past studies in comparative judgments have utilized different stimuli and studied comparative judgments across a range of dimensions such as magnitudes of digits, size of objects, and pleasantness of stimuli. These studies have included comparisons of numerical information (Foltz et al., 1984: Jamieson and Petrusic, 1975) as well as verbal information (Holyoak and Walker, 1976; Parkman, 1971). Two important effects, referred to as the symbolic distance effect and the semantic congruity effect (cf. Banks and Flora 1977), have consistently been observed in these studies.

The <u>symbolic distance effect</u> is the finding that as the distance between two stimuli along a dimension increases, faster (or more accurate) comparisons between the two stimuli are made by subjects. For example, in a comparison task involving digits,

a comparison between '1' and '100' is made faster (and/or more accurately) than a comparison between '1' and '3'. Studies have demonstrated the symbolic distance effect for comparisons involving numerical as well as verbal labels (Banks et al. 1976).

The semantic congruity effect has been described by Jaffe-Katz et al. (1989) as the phenomenon where "comparisons are faster when the instructions are congruent (or 'match') the stimuli than when they 'disagree' with them." For example, if a task requires a choice of the 'larger' item on a dimension such as magnitude of digits, decisions are made faster by subjects for a pair of large stimuli (such as '101' and '99') than for a pair of small stimuli (such as '3' and '1'), even though the difference in magnitude between the stimuli is the same. The semantic congruity effect has been obtained in several studies involving comparisons of numerical as well as verbal labels (Banks et al. 1976; Jaffe-Katz et al. 1989).

One early attempt to explain the distance and congruity effects was by Holyoak (1978), who put forward the 'analog comparison' model. According to this model, subjects, when asked to compare two stimuli, compare the distance of both stimuli from a reference point to determine which one is nearer to the reference point. The reference point is implied in the question that is put forward by the experimenter (for example, if the question is 'which animal is smaller,' then the reference point for size is zero). Subjects repeatedly compare the distance of both stimuli from the reference point till they determine which

one is closer to the reference point. According to this model, larger distances between the two stimuli makes it easier for the subject to compare them with the reference point, thereby leading to the distance effect. Further, by invoking the Weber Law that smaller differences are more detectable when the magnitudes of the stimuli are smaller, Holyoak argued that as the two stimuli move closer to the reference point, they are compared more quickly.

Jaffe-Katz et al. (1989) studied comparisons of numerical and nonnumerical probability expressions. Specifically, they studied comparisons of expressions which were both verbal (VV), both numerical (NN), and one expression numerical and the other verbal (NV). They argued that nonnumerical probability expressions, due to their vagueness, tend to overlap more than numerical expressions, thereby necessitating more repeated observations to arrive at a comparative judgment. Hence, they hypothesized that nonnumerical expressions would take longer to compare than numerical expressions. They found in their experimental study that, as hypothesized, comparison times were greater for a verbal/verbal (VV) and a numerical/verbal (NV) pair compared to a numerical/numerical (NN) pair.

Past research on comparisons provides a basis to understand comparisons of brand attributes by consumers. In investigating comparisons of brands along attributes, it is important to note differences between brand attributes and the dimensions typically used in past research. The typical dimensions used in past

studies have been digits, size of animals, probability expressions, etc., where subjects did not have clear preferences for either the larger or the smaller stimuli. In natural consumer settings, where brand attributes have utility properties, subjects have clear preferences for levels of brand attributes. For example, in the case of calculators, subjects may have a preference for low values if the attribute is weight of the calculator. On the other hand subjects may have a preference for high values if the attribute is display width of the calculator. Thus, the utility properties of the attributes may result in the existence of natural reference points for that attribute, such as the 'ideal value' of that attribute.

The objectives of this study are first, to examine using response times, whether comparisons of numerical attribute values (NN) are easier than comparisons of verbal (VV) or verbal/numerical (NV), second to assess whether the distance and congruity effects are observed for attribute labels, and third, to see how the utility property of the attribute affects the comparative judgments of the attributes. The first two objectives are explored in Study 1, while the third objective is explored in Study 2. The hypotheses for Study 1 are formally stated below.

Hypothesis 1 The response times for comparative judgments of attribute labels will be lesser for a pair of numerical labels (NN), than for a pair of verbal labels (VV) or a pair of verbal/numerical labels (NV).

Hypothesis 2 The symbolic distance effect will be observed for comparisons of labels of brand attributes.

Hypothesis 3 The semantic congruity effect will be observed for comparisons of labels of brand attributes.

EXPERIMENT 1

Method

Overview

A within-subject manipulation of mode (NN, NV, and NN), instructions ('choose higher' and 'choose lower'), and distance between stimuli was used in the study. Comparisons based on three combinations of stimuli (NN, NV, and NN) were used to test hypothesis 1. The distance between labels was manipulated to test the hypothesis on distance effect (hypothesis 2). Finally, the instructions ('choose higher' and 'choose lower') were also manipulated to test hypothesis 3 about the semantic congruity effect.

Materials

Subjects were required to perform comparisons along a single product attribute. The attribute chosen was display width of calculators. Labels used for this attribute were determined on the basis of a pretest. The pretest employed a magnitude scaling procedure and required subjects to estimate the magnitudes represented by a range of 13 verbal and 13 numerical labels for display width of calculators by drawing lines (or producing numbers) such that the length of the line (or the magnitude of the numbers) varied with the magnitude that was represented by the label (see Viswanathan and Childers 1992 for more details on pretesting). The pretesting procedure was used to determine the

number of levels of magnitude that were used (by identifying clusters of verbal labels) as well as equivalent numerical and verbal labels (by plotting estimated magnitudes against chosen verbal labels and identifying equivalent numerical labels). A three (mode; NN, NV, and VV) by two (instruction; choose higher versus choose lower attribute values) by four (distance between pairs; 1, 2, 3 and 4 units) within-subject design was employed in the study.

All possible combinations of the five numerical and verbal labels of the attribute, display width, were used in the study resulting in a total of 45 pairs of labels; 10 NN pairs, 10 VV pairs, and 25 NV pairs.

Procedure

Thirty two undergraduate students at a midwestern university participated in the study for course credit. The experiment was conducted using Macintosh computers. Subjects were familiarized with the set of labels used in the study and provided with instructions to perform the comparison as quickly as possible without compromising on accuracy.

The experimental stimuli was then administered. One pair of labels was presented at each trial. Subjects were required to indicate their response by clicking a mouse on buttons below each label. Each trial was separated by a masked screen for three seconds to mark the end of a trial. The order of trials was randomized across all subjects. Each subject performed comparison for two sets of forty five trials each, one involving

the 'choose higher' decision, and the other involving the 'choose lower' decision. The order of these two sets of trials was counterbalanced across two groups consisting of approximately equal number of subjects.

Results

The accurate responses were identified for the NN and VV conditions since these two conditions had objective criteria for determining accuracy. The mean response times for accurate responses for NN and VV, and mean response times for NV were computed for each subject for each level of distance and instruction. Several ANOVAs were run on the data and are described below.

Speed of Comparison (Hypothesis 1)

A two (instruction; choose higher versus choose lower) by three (mode; NN, NV, and VV) factorial ANOVA led to a significant main effect for mode ($F(2,62) = 19.41$; $p<0.001$), and a significant interaction between instruction and mode ($F(2,62) = 4.12$; $p<0.05$). The mean response time for the NN condition was found to be significantly less than those for NV ($F(1,62) = 33.42$; $p<0.001$) and VV ($F(1,62) = 24.04$; $p<0.001$) conditions, with the difference between NV and VV being non-significant ($F(1,62) = 0.77$; $p>0.35$). Means for NN, NV, and VV were 3.87, 4.34, and 4.27s respectively. These results support hypothesis 1. An examination of the interaction between instruction and mode suggested that the mean response time for 'choose higher' condition was significantly less than the 'choose lower'

condition ($F(1,31) = 5.69$; $p<0.05$), when both stimuli were verbal (VV). Also, the 'choose higher' decisions were made faster than 'choose lower' decisions, though the difference was not significant.

Symbolic Distance Effect (Hypothesis 2)

A two (instruction) by two (mode) by four (distance) factorial ANOVA led to a significant main effect for distance ($F(3,93) = 7.11$; $p<0.001$), a significant interaction between mode and distance ($F(6,186) = 2.52$; $p<0.05$), and a significant three-way interaction ($F(6,186) = 2.69$; $p<0.05$). An examination of the interaction between distance and mode revealed significant distance effects for NV ($F(3,93) = 7.27$; $p<0.001$) and VV ($F(3,93) = 5.7$; $P<0.001$), but not for NN ($F(3,93 = 1.64$; $p>0.18$). An examination by task revealed a significant distance effect for the 'choose higher' instruction ($F(3,93) = 7.36$; $p<0.001$), but not for the 'choose lower' instruction. On visual examination, none of the trends was monotonically decreasing except for VV in the 'choose higher' condition.

Linear trend analyses were performed to test for distance effects at each level of mode and at each level of instruction. They revealed significant linear trends for NV for 'choose higher' instruction ($F(1,93) = 7.77$; $p<0.01$), and VV for 'choose higher' instruction ($F(1,93) = 13.32$; $p<0.001$). Thus, hypothesis 2 was supported for the 'choose higher' instructions, for both VV and NV conditions, thereby suggesting that the distance effect may exist for these conditions.

Semantic Congruity Effect (Hypothesis 3)

The semantic congruity effect relates to an interaction between the size or magnitude of labels being compared and the instructions. Therefore, labels that were both high or both low were selected for further analysis. A two (instruction) by three (mode) by two (size; both labels high, both labels low) factorial ANOVA led to a significant interaction between instruction and size ($F(1,31) = 6.02$; $p<0.05$). 'Choose higher' judgments were made faster than 'choose lower' when both labels were high, whereas 'choose lower' judgments were made faster when both labels were low (Figure 1). Thus, hypothesis 3 was supported and an effect similar to the semantic congruity effect was found here.[1]

Insert Figure 1 about here

Discussion

Several interesting findings emerged from the analysis. First, the lower response times for NN labels indicated that they may be easier to compare than NV or VV labels. Further, the distance effect was not obtained for NN. This is an interesting finding since the presence of the distance effect would imply that distances between attribute values are utilized by subjects

---

[1]These analyses were repeated after deletion of outliers. Quantitatively similar results were obtained. One exception was that a significant linear trend was obtained for NN as well as 'choose higher' judgment.

in making comparisons. An absence of the distance effect would, therefore, imply that comparison are made without taking into consideration the distances between the stimuli. Therefore, the lack of a distance effect for NN suggests that a subjective interpretation of magnitudes based on distance may have been occurred to a lesser degree for NN when compared to NV or VV. Finally, the distance effect was found for 'choose higher' judgments, but not for 'choose lower' judgments, thereby suggesting greater subjective interpretation of magnitudes for the 'choose higher' judgments. Also, the interaction between instruction and size provided support for the semantic congruity effect.

It was discussed earlier that one important difference between the brand attributes and the stimuli used in other studies is that brand attributes have utility properties attached to them. Subjects may have made faster comparisons for the 'choose higher' condition and not for the 'choose lower' condition since the comparison in the 'choose higher' condition was directionally consistent with the utility properties of display width (on the assumption that utility for display width increases with increasing display width). The results obtained suggest a possible interaction between the vector properties of the attribute (increasing or decreasing) with task ('choose higher' or 'choose lower'). This possible interaction can be explained using Holyoak's (1978) reference point model. Holyoak's model assumed that the reference point was provided by

the instructions posed to subjects. However, if subjects already possess a reference point (such as an 'ideal point'), the correspondence between the existing reference point and the one provided by the instructions may be a factor in influencing the ease of comparisons. If the reference point provided in the instructions corresponds to subjects' ideal point, comparisons may be facilitated. On the other hand, if the reference point provided in the instructions is different from subject's ideal point, then comparisons may take more time.

If the explanation provided above is true, then an interaction should exist between the vector property of the attribute and the task. To see whether this interaction effect existed, another experiment was performed where the attribute of interest had a negative utility property, i.e., as the attribute value increased, its utility decreased. More formally, the following hypothesis was tested.

Hypothesis 4 For attributes with positive utility properties, the comparisons will be faster for the 'choose higher' instruction than the 'choose lower' instruction. For attributes with negative utility properties comparisons will be faster for the 'choose lower' instruction than the 'choose higher' instruction.

EXPERIMENT 2

Method

Procedure

The basic procedure was similar to that described in the earlier experiment, except that the attribute, weight of the calculator, was chosen to represent an attribute with a negative

utility. Twenty four undergraduates participated in the study for course credit. The results are described below.

Results

Speed of Comparison (Hypothesis 1)

A two (task; 'choose higher' versus 'choose lower') by two (mode; NN, NV, VV) factorial ANOVA was performed on the reaction times. Again, a significant main effect was found for mode ($F(2,46) = 41.13$; $p<0.001$). The mean response time for the NN condition was significantly less than that for the VV ($F(1,46) = 24.44$; $p<0.01$) and the NV ($F(1,46) = 16.92$; $p<0.001$) conditions, thereby replicating the results of Experiment 1 and providing support for hypothesis 1. The mean response time for the VV condition was significantly less than that for the NV condition ($F(1,46) = 16.92$; $p<0.10$), in contrast to Experiment 1. The mean reaction times for the NN, NV and VV conditions were 3.67s, 4.45s, and 4.09s respectively. Also, the mean response time for the 'choose lower' task was marginally lower than that for the 'choose higher' condition ($F(1,23) = 3.12$; $p<0.10$). Moreover, there was a significant interaction effect between task and mode. The mean response times for the 'choose lower' task was significantly less than that for the 'choose higher' task only for the NN condition. This is in contrast to the results of the previous experiment where, directionally, it was found that the response time for the 'choose lower' task was higher than that for the 'choose higher' task. This difference is important since it suggests that the utility property of the attribute may

influence the effects found.

Symbolic Distance Effect (Hypothesis 2)

A two (task) by three (mode) by four (distance) factorial ANOVA did not show a significant distance main effect ($F(3,69) = 1.83$; $p<0.15$). However, there was a significant interaction between mode and distance ($F(6,138) = 3.38$; $p<0.05$). A significant distance effect was found for the 'choose lower' task ($F(3,69) = 4.38$; $p<0.01$) and a marginally significant distance effect was found in the 'choose higher' task condition ($F(3,69) = 2.35$; $p<0.09$). A significant distance effect was found for the VV condition ($F(3,69) = 3.57$; $p<0.05$), and a marginally significant distance effect was found for NN ($F(3,69) = 2.29$; $p<0.09$). However, no significant effect was found for NV ($F(3,69) = 2.01$; $p<0.13$). Further, none of the distance effects for the various levels of mode appeared to be monotonically decreasing for increasing distance.

Semantic Congruity Effect (Hypothesis 3)

A two (instruction) by three (mode) by two (size; both labels high, both labels low) factorial ANOVA led to a marginally significant interaction between task and size ($F(1,23) = 3.03$; $p<0.10$). An examination showed that the mean response times for the 'choose lower' task was less than that for the 'choose higher' task, when both labels were low. Thus, the findings replicate the earlier findings and provide marginal support for the semantic congruity effect, thereby supporting hypothesis 3.

Combined Analysis (Hypothesis 4)

The data from the two experiments were combined in the next stage of analysis. A two (attribute type; positive versus negative utility property) by two (task) by three (mode) factorial between subjects ANOVA led to a significant interaction between attribute and task ($F(1,54) = 4.29$; $p<0.05$). For the attribute with the positive utility property, faster judgments were made in the 'choose high' condition compared to the 'choose low' condition, whereas, for the attribute with the negative utility property, faster judgments were made in the 'choose low' condition compared to the 'choose high' condition, thereby providing support for hypothesis 4 (see Figure 2). In addition, a significant interaction was also obtained between attribute type and mode ($F(2,108) = 3.97$; $p<0.05$). While for the NN and the VV conditions, the mean response times for judging the positive utility attribute was higher, for the NV condition, the mean response time for judging the negative utility attribute was higher.

Insert Figure 2 about here

Discussion

The second experiment shows that NN comparisons are faster than NV or VV comparisons. Also, the semantic congruity effect was found for attribute comparisons. In addition, the experiment provided evidence that the symbolic distance effect is mediated

by the utility properties of the attribute. For attributes with positive utility properties, the symbolic distance effect was found for the 'choose higher' task, and for attributes with negative utility properties, the symbolic distance effect was found for the 'choose lower' task. This is an important finding that has not been demonstrated in earlier research.

GENERAL DISCUSSION

This paper examined comparative judgments in consumer decision making. Using a comparative judgment task from cognitive psychology, comparisons of pairs of numerical and/or verbal brand attribute labels were studied in two experiments. In the first experiment, comparisons involving a pair of numerical labels were found to take less time than comparisons involving a pair of verbal or a numerical/verbal pair. This result suggests that comparisons involving a pair of numerical labels may be easier than involving a pair of verbal or a numerical/verbal pair, probably due to the more precise nature of the information in the case of numerical labels compared to the verbal labels.

In addition, several findings from cognitive psychology were tested in the context of comparison of product attributes. The results showed evidence for the existence of the semantic congruity effect, thereby suggesting that individuals access the meaning of a magnitude while making comparisons. The results did not provide consistent support for the symbolic distance effect. The results suggested that an important factor affecting

comparisons of numerical and verbal labels may be utility property of the attribute for which comparisons were being made. It appears that an implicit reference point may be provided by the ideal point on the attribute.

The research contributes to consumer research in several ways. First, it provides insights into the process of comparative judgments by studying the impact of various factors such as the nature of the reference points, the nature of instructions, and the distance between the labels on the ease of comparative judgments. Second, it looks at the impact of the information mode on the ease of comparative judgments.

Several avenues of future research are suggested here. One line of research can study the impact of implicit and explicit reference points in making comparisons. For example, the effect of product familiarity which may result in prior knowledge of mean value or the range of values on an attribute on comparative judgments can be examined. Another line of research can further focus on the distance and the congruity effects for comparisons of labels presented in different forms (for example, nutritional information preprocessed to different degrees and presented either as raw information or in relation to some standard, such as USRDA). In closing, it should be noted that the study of comparative judgments, apart from being of importance in itself, can also provide valuable insights about the nature of processing involved in decision making.

References

Banks, W.P., and J. Flora (1977). "Semantic and perceptual processes in symbolic comparisons." Journal of Experimental Psychology: Human Perception and Performance, 3, 278-290.

Banks, W.P., M. Fuji, and F. Kayra-Stuart (1976), "Semantic congruity effects in comparative judgments of magnitude of digits," Journal of Experimental Psychology: Human Perception & Performance, 2, 435-447.

Foltz, G., S. Poltrock, and G. Potts (1984). "Mental comparison of size and magnitude: Size congruity effects," Journal of Experimental Psychology: Learning, Memory & Cognition, 10, 442-453.

Jamieson, D.G., and W.M. Petrusic (1975). "Relational judgments with remembered stimuli," Perceptions and Psychophysics, 18, 373-378.

Holyoak, K.J. (1978). "Comparative judgments with numerical reference points," Cognitive Psychology, 10, 203-243.

Holyoak, K.J., and J.H. Walker (1976). "Subjective magnitude information in semantic orderings," Journal of Verbal Learning & Verbal Behavior, 15, 287-299.

Jaffe-Katz, A., D. Budescu, and T.S. Wallsten (1989). "Timed magnitude comparisons of numerical and nonnumerical expressions of uncertainty." Memory and Cognition, 17 (3), 249-264.

Moyer, R.S., and T.K. Landauer (1967). "Time required for judgments of numerical inequality," Nature, 215, 1519-1520.

Parkman, J.M. (1971), "Temporal aspects of digit and letter inequality judgments," Journal of Experimental Psychology, 91, 191-205.

Viswanathan, M. and T. Childers (1992), "The Encoding and Utilization of Magnitudes of Product Attributes: An Investigation Using Numerical and Verbal Information," Working Paper, University of Illinois.

Viswananthan, M. and S. Narayanan (1992), "The Processing of Numerical Versus Verbal Attribute Information: A Study Using Information Acquisition Patterns," to appear in Marketing Letters, 3:2.

FIGURE 1

RESULTS SHOWING THE SEMANTIC CONGRUITY EFFECT

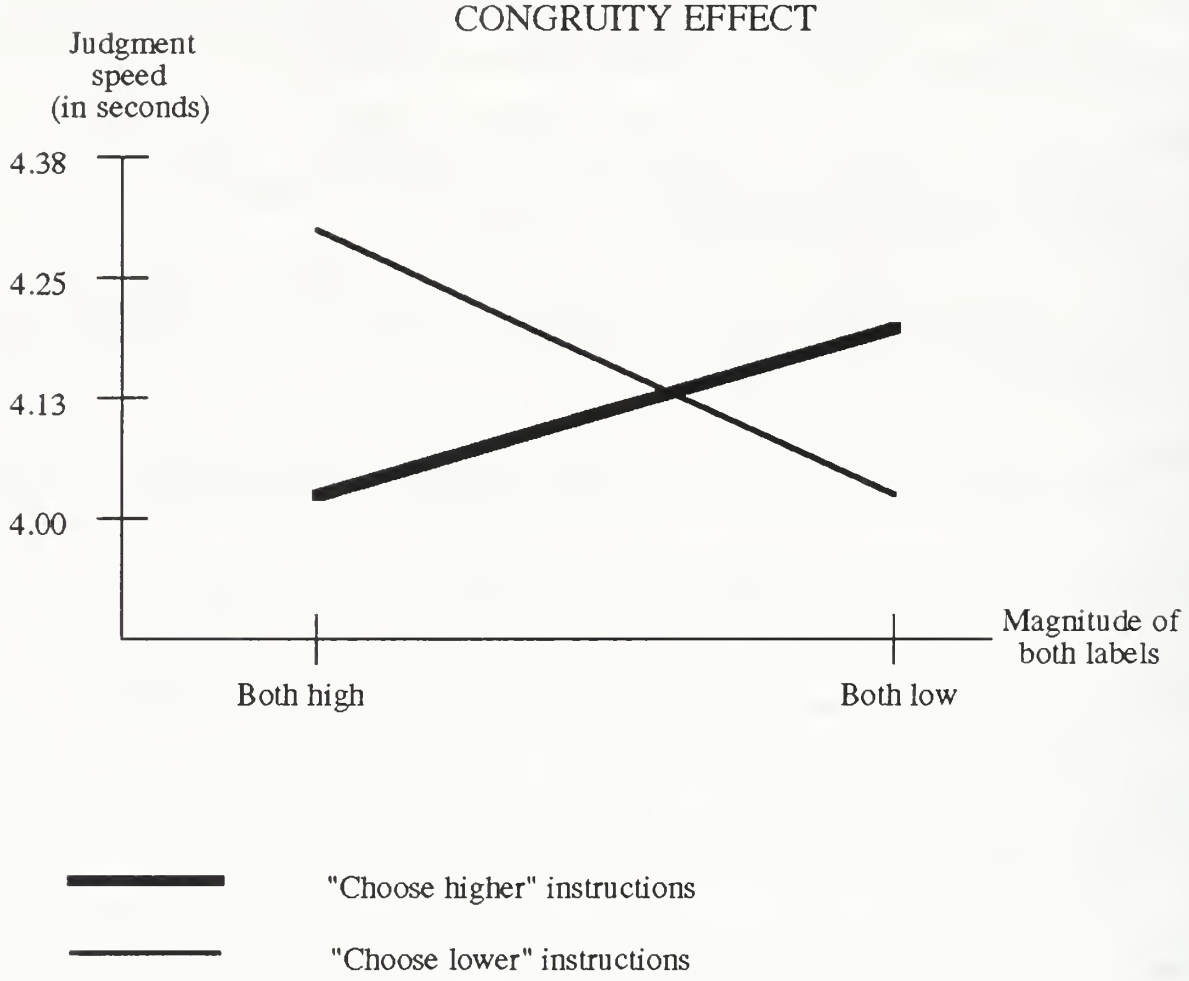

FIGURE 2

RESULTS SHOWING THE EFFECT OF UTILITY PROPERTIES OF AN ATTRIBUTE